

SEMANTIX ACADEMY



**PROJETO FINAL BIG DATA SCIENCE
ANÁLISE DOS MICRODADOS DO ENEM 2021**

Karen Dantas C.

SEMANTIX
2022

SUMÁRIO

1. INTRODUÇÃO.....	3
1.1. Contextualização	3
2. ANÁLISE DOS DADOS	4
2.1. Limpeza dos dados.....	4
2.2. Análise das características gerais da amostra	5
2.3. Análise geral das notas	8
2.4. Análise comparativa das notas considerando origem de escola pública ou privada .	10
2.4.1. Prova de Ciências da Natureza e suas tecnologias.....	11
2.4.2. Prova de Ciências Humanas e suas tecnologias	12
2.4.3. Prova de Linguagens, Códigos e suas tecnologias.....	13
2.4.4. Prova de Matemática e suas tecnologias	14
2.4.5. Prova de Redação	15
3. CONCLUSÕES	16
REFERÊNCIAS.....	17

1. INTRODUÇÃO

Este relatório tem por objetivo a análise dos dados coletados do ENEM (Exame Nacional do Ensino Médio) de 2021 sob a ótica da estatística descritiva, tendo como foco a comparação do desempenho entre alunos oriundos de escola pública e privada.

A pergunta a qual busca-se responder é: as pessoas que fizeram o ensino médio em escola pública tiveram desempenho inferior no ENEM de 2021 se comparadas às pessoas que estudaram em escola privada?

1.1. Contextualização

O ENEM é uma avaliação aplicada no Brasil desde 1998 que visa verificar se os participantes demonstram, “ao final do ensino médio, individualmente, domínio dos princípios científicos e tecnológicos que presidem a produção moderna e se detêm conhecimento das formas contemporâneas de linguagem” (INEP, 2022). A edição de 2021 contou com uma redação e 180 questões objetivas de múltipla escolha que foram aplicadas em dois dias no formato presencial e online. Segundo o INEP (2022), no primeiro dia foram realizadas as provas: Linguagens, Códigos e suas tecnologias; Redação; Ciências Humanas e suas tecnologias. E, no segundo dia, as provas: Ciências da Natureza e suas tecnologias; Matemática e suas tecnologias.

Garcia, Caldas e Torres (2021) criticaram o ENEM por ser uma avaliação padronizada, pois ignora o fato de a escola pública conter pessoas diversas e contextos plurais, e, conseqüentemente, é excludente, homogeneiza os sujeitos e é de cunho meritocrático. Garcia, Caldas e Torres (2021) ainda afirmam que os resultados do ensino da escola pública no ENEM transferem para o indivíduo uma responsabilidade que deveria ser do Estado, pois desconsidera a situação precária em que, muitas das vezes ocorre o processo educacional.

Além disso, o objetivo inicial do ENEM era “contribuir para a qualidade do Ensino Médio, e com isso induzir as mudanças no currículo, a partir da contrarreforma, percebe-se que ocorre a inversão desse objetivo” (GARCIA; CALDAS; TORRES, 2021). Por contrarreforma, as autoras se referem à reforma do ensino médio de 2017.

Jaloto e Primi (2021), afirmaram haver uma persistente desigualdade educacional no Enem. Para contornar isso, segundo, políticas públicas têm sido implementadas com o objetivo de superar as desigualdades educacionais tais como a Lei de Cotas de 2012 a qual visa garantir

a reserva de 50% das vagas de cursos do ensino superior das instituições federais para estudantes que cursaram integralmente o ensino médio em escolas públicas (BRASIL, 2012).

Porém, de acordo com Machado (2021), as condições de preparação para o ENEM são desiguais, e o aumento da competição para ingressar no ensino superior torna-se mais para alunos com menos recursos. Sendo assim, “melhorar as condições de acesso ao ensino superior envolve, inevitavelmente, melhorar as condições de aprendizado no Ensino Médio” (MACHADO, 2021).

Segundo Garcia, Caldas e Torres (2021), desde o início a intenção do MEC (Ministério da Educação) era utilizar o ENEM como principal mecanismo de seleção para o ensino superior.

Em relação ao ENEM de 2021, de acordo com Machado (2021), foi a edição de menor número de inscritos desde 2005. Uma das possíveis causas apontadas foi a pandemia de COVID-19 que se iniciou em 2020 e obrigou, durante um período, as escolas a adotarem o ensino remoto o que gerou muitas dificuldades para os alunos de diferentes contextos socioeconômicos. Entretanto, Machado (2021) afirma que o número vem caindo bem antes da pandemia o que sugere haver outras causas também.

2. ANÁLISE DOS DADOS

Os microdados estudados foram obtidos diretamente de um arquivo *.csv* disponível no site do INEP, e as análises estatísticas foram realizadas com o uso da ferramenta *Jupyter Notebook* e linguagem de programação *Python*¹ versão 3.

Além das funcionalidades padrões do *Python*, também foram utilizadas as bibliotecas *Pandas*² e *Seaborn*³.

2.1. Limpeza dos dados

Antes de se realizar a análise dos dados, primeiramente foi necessário realizar uma limpeza nos mesmos. As limpezas executadas foram:

- Eliminados todos os registros que possuíam valores nulos (1.151.725 registros), os quais representavam pessoas que não fizeram uma ou mais provas;

¹ Site oficial: <https://www.python.org/>

² Site oficial: <https://pandas.pydata.org/>

³ Site oficial: <https://seaborn.pydata.org/>

- Seleccionadas apenas as pessoas que já concluíram o ensino médio, ou que estavam cursando o ensino médio e concluiriam em 2021 (373.883 registros). Pois, assim seria possível avaliar efetivamente o ensino médio brasileiro;
- Excluídas as pessoas que não informaram se cursaram o ensino médio em escola particular ou pública (1.016.827 registros).

Após a limpeza, restaram 847.397 registros os quais, considerando uma população de pessoas que se inscreveram no ENEM 2021, representam uma amostra de pessoas que realizaram todas as provas do ENEM em 2021, são de escola pública ou privada, e que concluíram o ensino médio até o final 2021.

2.2. Análise das características gerais da amostra

A frequência relativa das pessoas que fizeram o ENEM 2021 e que informaram ter cursado o ensino médio em escola pública é de 78,83%, e de escola privada é 21,17%. O que pode indicar um maior interesse dos alunos de escola pública em cursar o ensino superior em uma instituição federal ou obter uma bolsa de estudos.

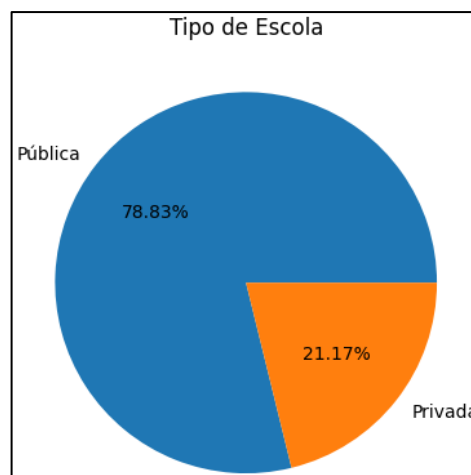


Figura 1 – Frequência relativa do tipo de escola.
Fonte: Elaborada pela autora.

Conforme pode ser observado na Figura 2, 60,17% dos participantes são do sexo feminino, e 39,83% do sexo masculino.

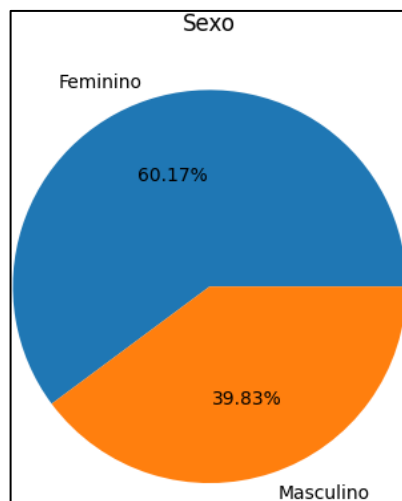


Figura 2 – Frequência relativa do sexo dos participantes.
Fonte: Elaborada pela autora.

Em relação à cor/raça (Figura3), mais da metade dos participantes se declararam como branco (45,30%) e pardo (39,98%). E a cor/raça com menor representatividade foi a indígena (0,48%). 2,10% não declararam sua cor/raça.

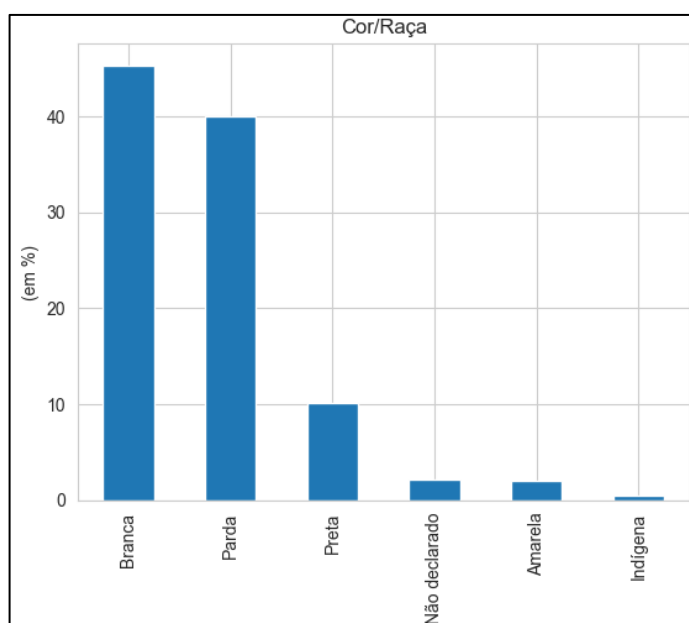


Figura 3 – Frequência relativa da cor/raça.
Fonte: Elaborada pela autora.

Como pode ser visto na Figura 4, no que diz respeito à UF (Unidade Federativa) da escola onde foi realizada a prova, a que teve maior frequência relativa foi a de São Paulo (SP) com 17,35% e a de menor frequência relativa foi Roraima (RR) com 0,19%.

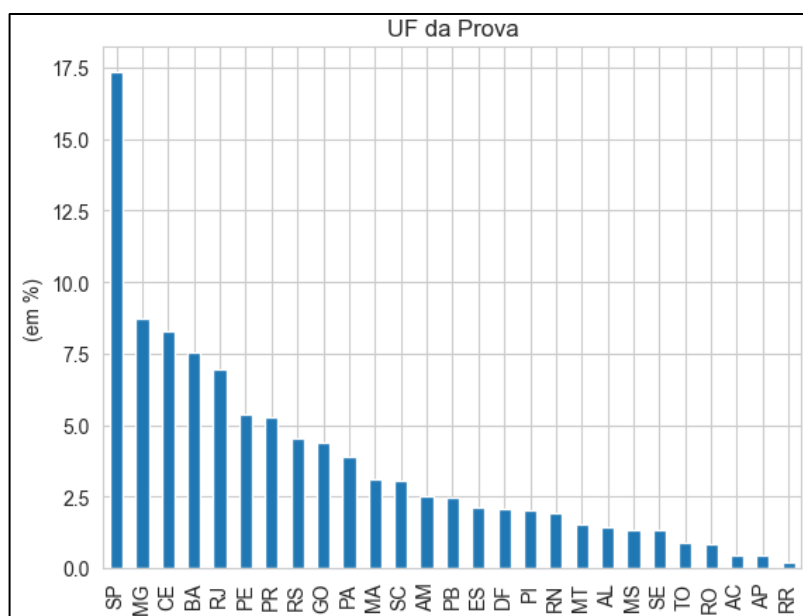


Figura 4 – Frequência relativa da UF da prova.
Fonte: Elaborada pela autora.

De acordo com a Tabela 1, mais de 85% da amostra de participantes é representada por pessoas de 17 anos (44,83%) e 18 anos (42,48%), e apenas 0,001% e 0,002% representam, respectivamente, pessoas de entre 66 e 70 anos, e maiores de 70 anos.

Tabela 1 - Frequência relativa da faixa etária.

Faixa etária	Frequência relativa (%)
17 anos	44,826
18 anos	42,479
19 anos	7,492
20 anos	2,100
Menor de 17 anos	1,186
21 anos	0,641
22 anos	0,265
Entre 26 e 30 anos	0,237
Entre 31 e 35 anos	0,147
23 anos	0,142
Entre 36 e 40 anos	0,110
24 anos	0,104
Entre 41 e 45 anos	0,076
25 anos	0,074
Entre 46 e 50 anos	0,056
Entre 51 e 55 anos	0,035
Entre 56 e 60 anos	0,019
Entre 61 e 65 anos	0,007
Entre 66 e 70 anos	0,002
Maior de 70 anos	0,001

Fonte: Elaborada pela autora.

E, por fim, no que tange a nacionalidade 97,55% são brasileiros, 0,13% estrangeiros e 0,11% não informaram a sua nacionalidade, conforme a Tabela 2.

Tabela 2 - Frequência relativa da nacionalidade.

Nacionalidade	Frequência relativa (%)
Brasileiro(a)	97,55
Brasileiro(a) Naturalizado(a)	1,97
Brasileiro(a) Nato(a), nascido(a) no exterior	0,24
Estrangeiro(a)	0,13
Não informado	0,11

Fonte: Elaborada pela autora.

2.3. Análise geral das notas

Considerando as 847.397 pessoas de escola pública ou privada que fizeram o ENEM 2021, na Tabela 3 e Figura 5 estão os dados da estatística descritiva das variáveis quantitativas referentes às notas das provas, onde: NU_NOTA_CN é referente as notas da prova de Ciências da Natureza e suas tecnologias (CN); NU_NOTA_CH é referente a Ciências Humanas e suas tecnologias (CH); NU_NOTA_LC é referente a Linguagens, Códigos e suas tecnologias (LC); NU_NOTA_MT é referente a Matemática e suas tecnologias (MT); e, por fim, NU_NOTA_REDACAO é referente a Redação.

No geral, a nota mínima de todas as provas foi 0 e a única prova que obteve a nota máxima de 1.000 pontos foi a de Redação. E, de acordo com o gráfico boxplot da Figura 5, desconsiderando-se os valores discrepantes (*outliers*), a nota mínima das provas CN, CH, LC e MT está em torno de 300 pontos, e a de Redação 200 pontos; já em relação a nota máxima, a nota das provas de CN e LC está em torno de 700 pontos, a de CH próxima a 800, a de MT 850 e a de Redação 1.000 pontos.

Em relação à média, as provas que mais obtiveram notas maiores que a metade do total de pontos foram a de Redação com 615,04 pontos (com desvio padrão de 189,22), a de MT com 529,69 pontos (com desvio padrão de 107,63) e CH com 513,07 pontos (com desvio padrão de 90,80). Esses resultados indicam que as escolas do ensino médio teriam que rever, primeiramente, o ensino do conteúdo das provas CN e LC, em especial a CN que teve o pior resultado.

Já em relação a mediana que é uma medida que é menos afetada pelos *outliers*, as variáveis que variaram menos em relação à média foram a CH e LC, com 513,1 e 500,7 pontos, respectivamente. E, embora a LC tenha passado dos 500 pontos, ela continua sendo a segunda prova de menor nota.

Além disso, pelo menos 25% das pessoas tiraram menos que 424,9 pontos na prova de CN, 441,9 nas provas de CH e MT, 444,8 na prova de LC e 520 na prova de Redação. E pelo menos 25% tiraram mais que 536,7 pontos na prova CN, 579,4 na CH, 552,2 na LC, 605,7 na MT e 740 na Redação.

De modo geral, analisando-se o terceiro quartil (75%), os resultados apontam que grande parte das pessoas tiraram notas distantes da nota máxima, o que indica que o ensino médio brasileiro e a prova do ENEM têm que ser revistos.

Tabela 3 – Notas de pessoas de escola pública ou privada.

	count	mean	std	min	25%	50%	75%	max
NU_NOTA_CN	847397.0	484.17	77.90	0.0	424.9	474.3	536.7	848.7
NU_NOTA_CH	847397.0	513.07	90.80	0.0	441.9	513.1	579.4	846.9
NU_NOTA_LC	847397.0	498.15	76.14	0.0	444.8	500.7	552.2	820.5
NU_NOTA_MT	847397.0	529.69	107.63	0.0	441.9	509.1	605.7	953.1
NU_NOTA_REDACAO	847397.0	615.04	189.22	0.0	520.0	600.0	740.0	1000.0

Fonte: Elaborada pela autora.

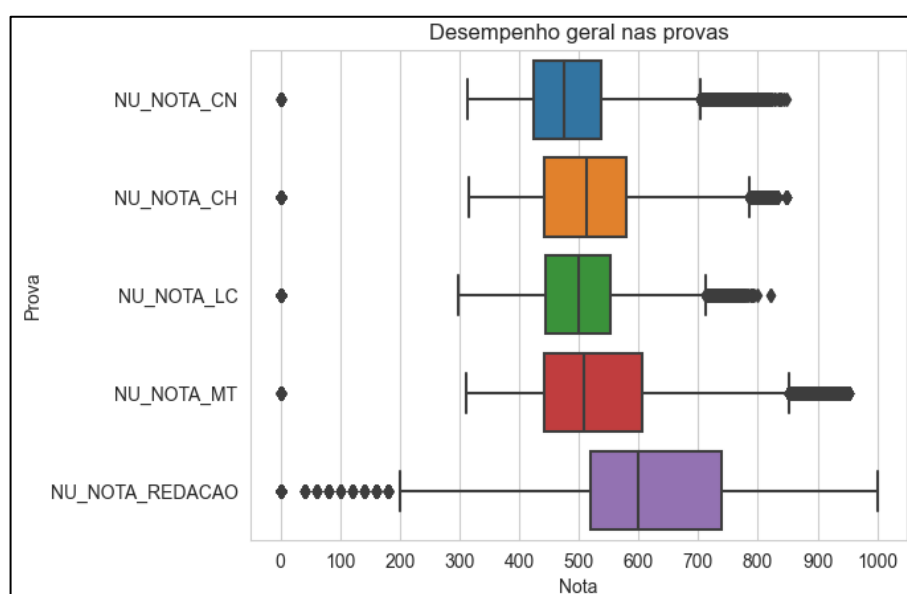


Figura 5 – Boxplot das notas em geral.

Fonte: Elaborada pela autora.

2.4. Análise comparativa das notas considerando origem de escola pública ou privada

Nas Tabela 4 e Tabela 5 estão os dados da estatística descritiva das variáveis quantitativas referentes às notas das provas das pessoas que cursaram o ensino médio em escola pública e privada, respectivamente. Do total de 847.397 pessoas, 668.036 estudaram em escola pública e 179.361 em escola privada.

Pode-se constatar de imediato que as pessoas de escola privada em todas as provas tiveram as médias e medianas de notas maiores do que as de escola pública. O corrobora com os resultados das edições do ENEM de 2018 e 2019. Jaloto e Primi (2021), ao analisarem os dados do ENEM 2018 constataram que ser de escola pública associou-se a uma redução no desempenho, em contrapartida ser de escola privada associou-se a um aumento. Além disso,

Souza (2021) analisou os dados do ENEM 2019, e concluiu que as pessoas oriundas de escolas particulares obtiveram resultados superiores em todas as provas.

Nos subtópicos a seguir são analisados os desempenhos em cada prova, considerando-se separadamente as pessoas que cursaram o ensino médio em escola pública das que estudaram em escola privada.

Tabela 4 – Notas de pessoas de escola pública.

	count	mean	std	min	25%	50%	75%	max
NU_NOTA_CN	668036.0	469.46	69.43	0.0	417.3	461.4	514.6	848.7
NU_NOTA_CH	668036.0	498.44	85.23	0.0	432.2	496.0	561.3	846.9
NU_NOTA_LC	668036.0	485.26	72.38	0.0	434.1	487.8	536.8	820.5
NU_NOTA_MT	668036.0	508.64	95.32	0.0	433.1	488.9	571.5	953.1
NU_NOTA_REDACAO	668036.0	583.28	185.20	0.0	500.0	580.0	680.0	1000.0

Fonte: Elaborada pela autora.

Tabela 5 – Notas de pessoas de escola privada.

	count	mean	std	min	25%	50%	75%	max
NU_NOTA_CN	179361.0	538.99	83.08	0.0	477.2	542.5	598.9	844.5
NU_NOTA_CH	179361.0	567.55	90.16	0.0	510.5	577.3	630.3	846.9
NU_NOTA_LC	179361.0	546.16	70.41	0.0	502.2	551.8	595.4	799.3
NU_NOTA_MT	179361.0	608.09	114.45	0.0	519.5	615.6	691.5	953.1
NU_NOTA_REDACAO	179361.0	733.34	153.80	0.0	620.0	740.0	860.0	1000.0

Fonte: Elaborada pela autora.

2.4.1. Prova de Ciências da Natureza e suas tecnologias

A Figura 6 contém o boxplot referente às notas da prova de CN.

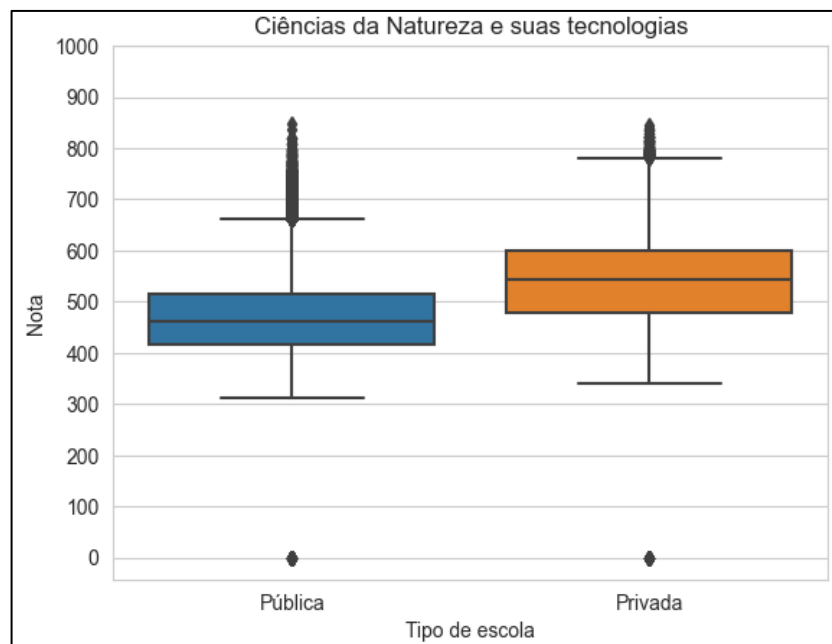


Figura 6 – Boxplot da prova de Ciências da Natureza e suas tecnologias.
Fonte: Elaborada pela autora.

A mediana das notas das pessoas de escola privada foi de 542,5 pontos que é 81,1 pontos superior aos 461,4 pontos das pessoas de escola pública. Desconsiderando-se os *outliers*, a nota mínima e máxima dos participantes da escola pública é, respectivamente, em torno de 300 e 660 pontos, e da escola privada é em torno de 350 e 790 pontos.

Ademais, pelo menos 25% das pessoas da escola pública tiraram menos que 417,3 pontos e 75% tiraram até no máximo 514,6 pontos. Já em relação à escola privada, pelo menos 25% tiraram menos que 477,2 pontos e 75% tiraram até no máximo 598,9 pontos.

A partir dos resultados obtidos, pode-se concluir que as pessoas que cursaram o ensino médio em escola privada tiveram melhor desempenho na prova de CN, em relação aos que cursaram em escola pública.

2.4.2. Prova de Ciências Humanas e suas tecnologias

Na Figura 7 está contido o boxplot referente às notas da prova de CH.

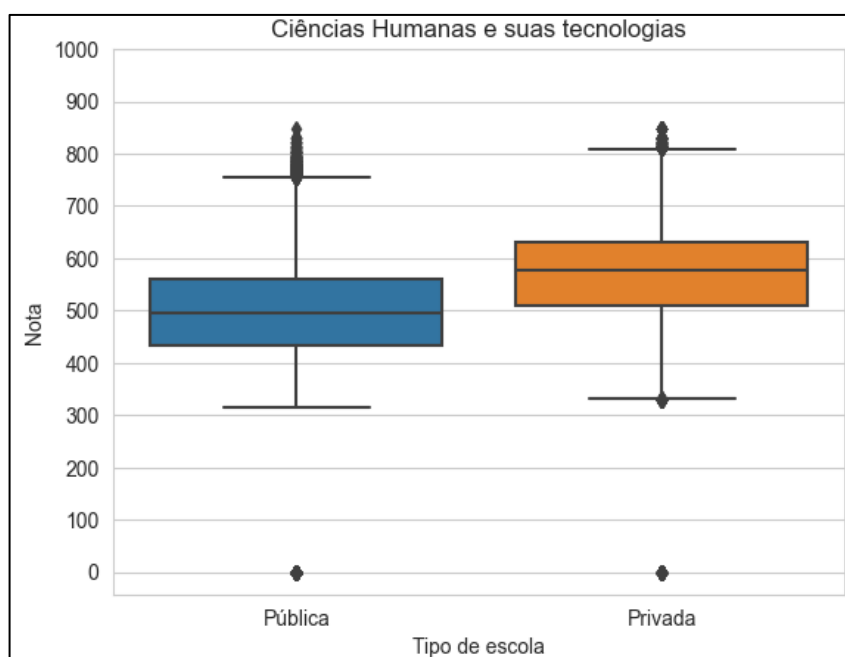


Figura 7 – Boxplot da prova de Ciências Humanas e suas tecnologias.
Fonte: Elaborada pela autora.

A mediana das notas das pessoas de escola privada foi de 577,3 pontos que é 81,3 pontos superior aos 496 pontos das pessoas de escola pública. Desconsiderando-se os *outliers*, a nota mínima e máxima dos participantes da escola pública é, respectivamente, em torno de 320 e 760 pontos, e da escola privada é em torno de 330 e 810 pontos.

Além disso, pelo menos 25% das pessoas da escola pública tiraram menos que 432,2 pontos e 75% tiraram até no máximo 561,3 pontos. Já em relação à escola privada, pelo menos 25% tiraram menos que 510,5 pontos e 75% tiraram até no máximo 630,3 pontos.

A partir dos resultados obtidos, pode-se inferir que as pessoas que cursaram o ensino médio em escola privada também tiveram melhor desempenho na prova de CH.

2.4.3. Prova de Linguagens, Códigos e suas tecnologias

Na Figura 8 está representado o boxplot referente às notas da prova de LC.

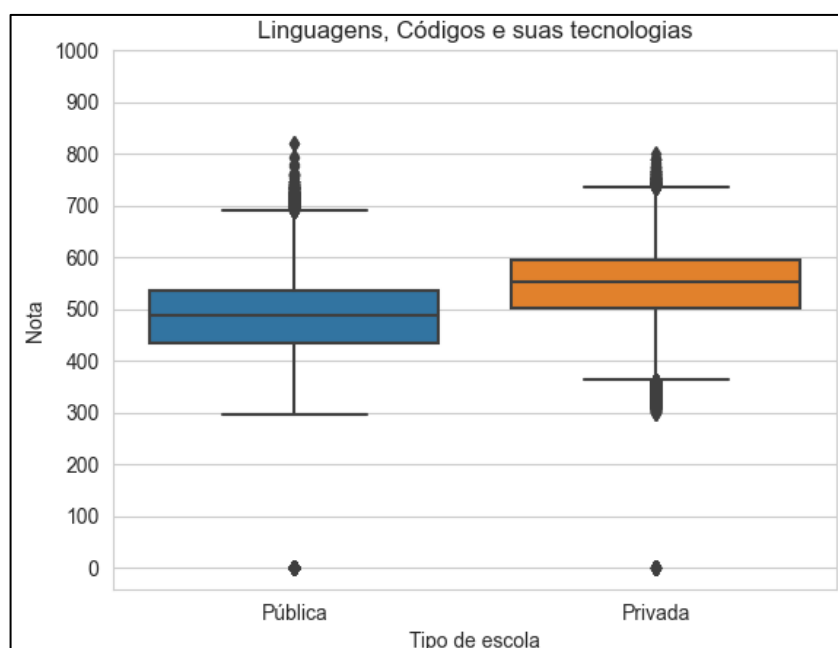


Figura 8 – Boxplot da prova de Linguagens, Códigos e suas tecnologias.
Fonte: Elaborada pela autora.

A mediana das notas das pessoas de escola privada foi de 551,8 pontos que é 64 pontos superior aos 487,8 pontos das pessoas de escola pública. Desconsiderando-se os *outliers*, a nota mínima e máxima dos participantes da escola pública é, respectivamente, em torno de 300 e 700 pontos, e da escola privada é em torno de 360 e 740 pontos.

Ademais, pelo menos 25% das pessoas da escola pública tiraram menos que 434,1 pontos e 75% tiraram até no máximo 536,8 pontos. Já em relação à escola privada, pelo menos 25% tiraram menos que 502,2 pontos e 75% tiraram até no máximo 595,4 pontos.

Pode-se concluir que as pessoas que cursaram o ensino médio em escola privada também tiveram melhor desempenho na prova de LC, a partir dos resultados obtidos.

2.4.4. Prova de Matemática e suas tecnologias

A Figura 9 contém o boxplot referente às notas da prova de MT.

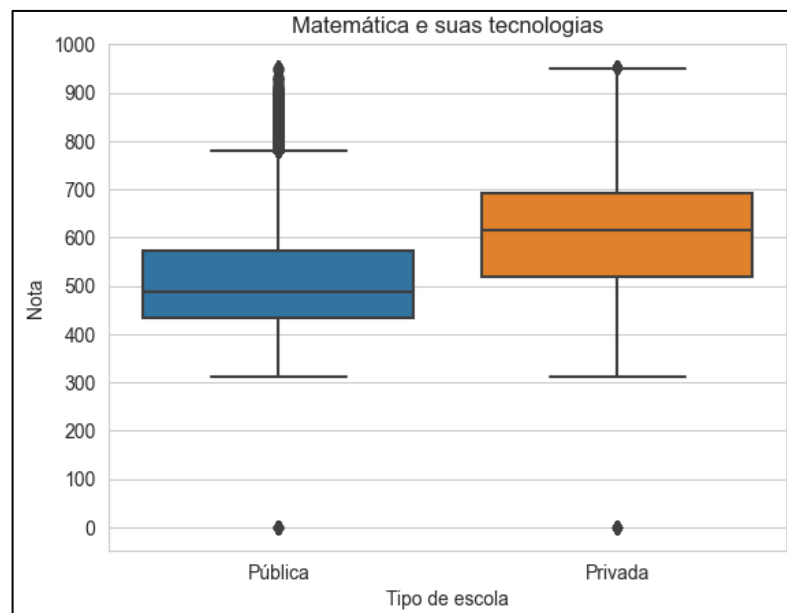


Figura 9 – Boxplot da prova de Matemática e suas tecnologias.
Fonte: Elaborada pela autora.

A mediana das notas das pessoas de escola privada foi de 615,6 pontos que é 126,7 pontos superior aos 488,9 pontos das pessoas de escola pública. Essa diferença de pontuação é bastante considerável.

Desconsiderando-se os *outliers*, a nota mínima e máxima dos participantes da escola pública é, respectivamente, em torno de 310 e 780 pontos, e da escola privada é em torno de 310 e 950 pontos.

Além disso, pelo menos 25% das pessoas da escola pública tiraram menos que 433,1 pontos e 75% tiraram até no máximo 571,5 pontos. Já em relação à escola privada, pelo menos 25% tiraram menos que 519,5 pontos e 75% tiraram até no máximo 691,5 pontos.

A partir dos resultados obtidos, pode-se inferir que as pessoas que cursaram o ensino médio em escola privada também tiveram melhor desempenho na prova de MT, com diferença de 126,7 pontos na mediana se comparado à escola pública.

2.4.5. Prova de Redação

Na Figura 10 está contido o boxplot referente às notas da prova de Redação.

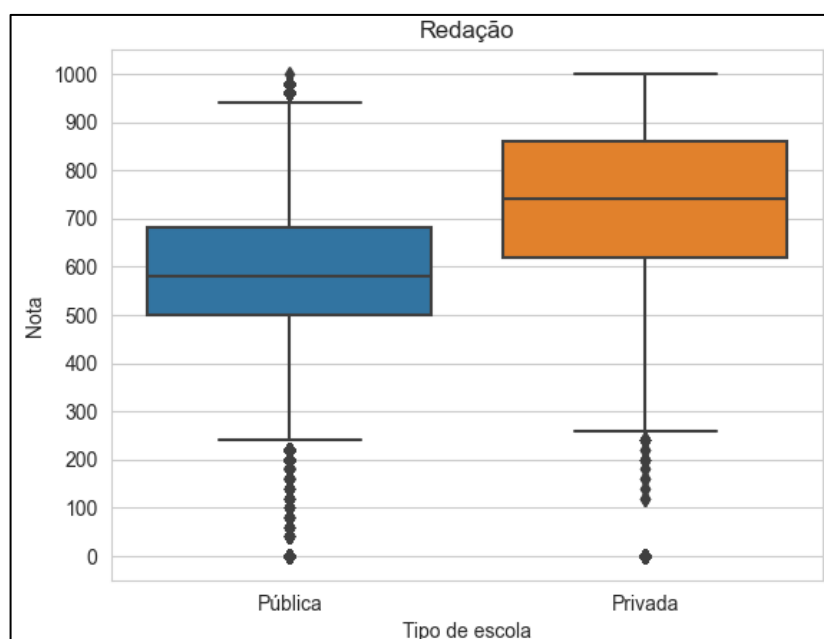


Figura 10 – Boxplot da prova de Redação.

Fonte: Elaborada pela autora.

A mediana das notas das pessoas de escola privada foi de 740 pontos que é 160 pontos superior aos 580 pontos das pessoas de escola pública. Essa diferença de pontuação é a maior obtida entre as análises realizadas.

Desconsiderando-se os *outliers*, a nota mínima e máxima dos participantes da escola pública é, respectivamente, em torno de 250 e 950 pontos, e da escola privada é em torno de 270 e 1.000 pontos.

Além disso, pelo menos 25% das pessoas da escola pública tiraram menos que 500 pontos e 75% tiraram até no máximo 680 pontos. Já em relação à escola privada, pelo menos 25% tiraram menos que 620 pontos e 75% tiraram até no máximo 860 pontos.

A partir dos resultados obtidos, pode-se concluir que as pessoas que cursaram o ensino médio em escola privada também tiveram melhor desempenho na prova de Redação, sendo que há uma diferença considerável de 160 pontos na mediana se comparado à escola pública.

3. CONCLUSÕES

O ENEM é uma avaliação anual a qual milhões de brasileiros fazem para avaliar seus conhecimentos ou ingressar no ensino superior.

Neste relatório foi analisada uma amostra de pessoas que realizaram todas as provas do ENEM em 2021, são de escola pública ou privada, e que concluíram o ensino médio até o final

2021, totalizando 847.397 registros sendo que 78,83% cursaram o ensino médio em escola pública e 21,17% em escola privada.

A partir das análises realizadas, pode-se concluir que as pessoas de escola privada obtiveram melhor desempenho em todas as provas do ENEM 2021 se comparados às pessoas oriundas da escola pública. Essa informação indica o ensino médio das escolas públicas deve ser revisto de forma que o conteúdo a ser lecionado seja no mínimo equivalente à da escola privada. Além disso, também pode-se inferir que os alunos de escola pública também estão em desvantagem quando forem concorrer a vagas no ensino superior através da nota no ENEM 2021, sendo que muitos alunos teriam que estudar previamente conteúdos extras para complementar o conhecimento.

Outros estudos realizados considerando edições do ENEM de 2018 e 2019 também concluíram que o desempenho das pessoas de escola pública foi inferior às de escola privada.

Ademais, os resultados também apontam que, independentemente do tipo de escola de origem, grande parte das pessoas tiraram notas distantes da nota máxima, o que indica que muitos dos conteúdos exigidos na prova não foram absorvidos pelos alunos, não foram lecionados no ensino médio ou não foram bem interpretados. É imprescindível que sejam feitos estudos que busquem identificar o que pode ser melhorado no ensino médio brasileiro e na prova do ENEM, de forma que uma educação de qualidade possa ser ofertada a todos e uma avaliação de conhecimento justa seja aplicada.

REFERÊNCIAS

BRASIL. **Lei nº 12.711**, de 29 de agosto de 2012. Dispõe sobre o ingresso nas universidades federais e nas instituições federais de ensino técnico de nível médio e dá outras providências. Diário Oficial da União, Brasília, DF, 30 ago. 2012. Seção 1, p. 1.

GARCIA, F. M.; CALDAS, R. S. M.; TORRES, G. C. O ENEM como política de avaliação e as contradições ao processo de democratização educacional. **Perspectiva**, Florianópolis, SC, v.39, n.3, p.1-21, 2021. Disponível em: < <https://doi.org/10.5007/2175-795X.2021.e68157> >. Acesso em: 03 dez. 2022.

INSTITUTO NACIONAL DE ESTUDOS E PESQUISAS EDUCACIONAIS ANÍSIO TEIXEIRA (INEP). Microdados do Enem 2021. Brasília: **Inep**, 2022. Disponível em: < <https://www.gov.br/inep/pt-br/acesso-a-informacao/dados-abertos/microdados/enem> >. Acesso em: 30 nov. 2022.

JALOTO, A. ; PRIMI, R. Fatores socioeconômicos associados ao desempenho no Enem. **Em Aberto**, Brasília, DF, v. 34, n. 112, p. 125-141, 2021. Disponível em:

<<http://rbep.inep.gov.br/ojs3/index.php/emaberto/article/view/5002>>. Acesso em: 03 dez. 2022.

MACHADO, C. Aumento da competição no Enem torna acesso mais difícil para alunos com menos recursos. **Folha de S.Paulo**, São Paulo SP, 22 de nov. de 2021. Disponível em: <<https://www1.folha.uol.com.br/colunas/cecilia-machado/2021/11/aumento-da-competicao-no-enem-torna-acesso-mais-dificil-para-alunos-com-menos-recursos.shtml>>. Acesso em: 03 dez. 2022.

SOUZA, T. O. de. **Análise de dados: um estudo do perfil dos participantes do ENEM 2019**. TCC (Graduação) - Curso de Ciência da Computação, Universidade Federal Rural do Semi-árido, Mossoró, RN, 2021. Disponível em: <<https://repositorio.ufersa.edu.br/handle/prefix/6916>>. Acesso em: 04 dez. 2022.